AD_____

Award Number: DAMD17-02-1-0489

TITLE: Computer-Aided Interval Change Analysis of Microcalcifications on Mammograms for Breast Cancer Detection

PRINCIPAL INVESTIGATOR: Lubomir Hadjiiski, Ph.D.

CONTRACTING ORGANIZATION: University of Michigan
Ann Arbor, Michigan 48109-1274

REPORT DATE: July 2003

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 2003	Annual (1 Jul 02-30 Jun 03)

4. TITLE AND SUBTITLE
Computer-Aided Interval Change Analysis of Microcalcifications on Mammograms for Breast Cancer Detection

5. FUNDING NUMBERS
DAMD17-02-1-0489

6. AUTHOR(S)
Lubomir Hadjiiski, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Michigan
Ann Arbor, Michigan 48109-1274

E-Mail: lhadjisk@umich.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
Original contains color plates. All DTIC reproductions will be in black and white.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 Words)

The goal of this project is to develop a computer-aided diagnosis (CAD) system for interval change analysis of lesions on mammograms. An important component of the CAD system is the multistage regional registration technique for identifying corresponding microcalcification clusters on temporal pairs of mammograms. In the first stage, an initial search region was estimated on the prior mammogram based on the lesion location on the current mammogram. In the second stage the search region was refined. In the third stage the lesion was detected within the search region. In the first stage we compared the regional registration method (RRM) to the use of linear and nonlinear warping techniques for the initial estimation of the lesion location. 390 temporal pairs of mammograms were used for evaluation. The average distance between the estimated and the true lesion centroids on the previous mammogram after the initial stage was 8.5±6.2mm for RRM and 9.0±6.7mm for the best of the warping techniques. The RRM method outperformed the warping techniques. In the second step, automated detection of microcalcification cluster within the search region is performed. Using our current cluster detection program with standard thresholds, 69.4% (50/72)TP with 0.21 false positives (FP) were detected within the search region. Using a high-sensitivity threshold, 84.7% (61/72) TP with 0.75 FP were detected.

14. SUBJECT TERMS
Breast Cancer, Computer-aided Diagnosis, Screening, Classification, Image Analysis

15. NUMBER OF PAGES
20

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

(3) Table of Contents

(4) Introduction

Treatment of breast cancer at an early stage can significantly improve the survival rate of patients. Mammography is currently the most sensitive method for detecting early breast cancer, and it is also the most practical for screening. Although general rules for differentiation between malignant and benign lesions exist, in clinical practice, approximately only 15-30% of cases referred to surgical biopsy are actually malignant. A number of research groups are in the process of developing computer-aided diagnosis (CAD) methods which can provide a consistent and reproducible second opinion to the radiologist for the detection and classification of breast abnormalities.

Radiologists routinely compare mammograms from a current examination with those obtained in previous years, if available, for identifying interval changes, detecting potential abnormalities, and in evaluating breast lesions. It is widely accepted that interval changes in mammographic features are very useful for both detection and classification of abnormalities. However, CAD techniques that use multiple exams for detection or characterization have not been commonly explored, probably because of the difficulty in the registration of the compressed breast images from different exams. We have been investigating methods for analysis of temporal changes of masses on mammograms to improve detection and classification. To our knowledge, there is no existing CAD technique for registration of microcalcification clusters or classification of microcalcifications based on temporal change information.

The extraction of any meaningful information from a prior mammogram first requires a common frame of reference between the current and prior mammograms. Several complicating factors, such as breast compression difference between current and prior mammograms, energy difference between the two imaging conditions, differences in screen film properties and film processing conditions, and potential changes in breast structures between the two images with patient age, make it difficult to obtain such a frame of reference. On breast images, there are no invariant landmarks (except for the nipple) that can serve as control points in conventional image registration methods to register the two mammograms. In this project, we propose to develop an innovative regional registration method that does not depend on specific control points. We will first approximately align the current and prior mammogram based on maximization of mutual information. Next, we will design a novel approach in which the computer emulates the radiologists' search method in finding corresponding lesions on mammograms. Automated search of microcalcification cluster within the search region on the prior mammogram will be performed. Our current automated microcalcification detection algorithm will provide a basis for this search. However, since the detection is limited to the small search region, the detection can be performed in high resolution and the algorithm parameters can be adjusted to improve the detection sensitivity of the very subtle clusters on the prior mammograms without excessive trade-off in increasing false-positives. A correspondence classifier will be developed to identify the matched pair of clusters on the two mammograms. The image features of the corresponding microcalcification clusters can then be automatically extracted and feature measures characterizing interval changes derived. A classification scheme to differentiate malignant and benign clusters using the interval change information will be developed. This computerized interval change analysis will be an important component of a CAD system for mammographic interpretation.

This project aims at developing a novel interval change analysis scheme to improve the accuracy of CAD. We will investigate the problem of classifying microcalcifications as malignant or benign based on temporal changes in mammographic features using a combination of computer vision, automated feature extraction, statistical classification, and artificial

intelligence techniques. We hypothesize that the use of temporal information would improve the ability of CAD to offer an accurate and objective second opinion to radiologists which, in turn, would increase the positive predictive value of mammography, reduce the number of benign biopsies, and hence reduce both cost and patient morbidity. If integrated in a complete CAD system, the algorithms to be developed in this project may also increase the efficacy of mammography for early detection of breast cancer.

(5) Body

In the first year (7/1/02-6/30/03) of this grant, we have performed the following studies:

(A) Database collection of malignant and benign breast microcalcification cases that have multiple examinations (Task 1)

We started collecting the data set for this study from the files of patients who had undergone biopsy at the University of Michigan. The mammograms are scanned and the images are saved in our storage device using automated graphic user interface developed in our laboratory. Additionally the film information is recorded in a Microsoft Access database. Temporal pairs of images were obtained. The current mammogram of each temporal pair exhibited a biopsy-proven mass. We scan both cranio-caudal and mediolateral-oblique views. The mammograms were digitized with a LUMISCAN 85 laser scanner at a pixel resolution of 0.05 mm x 0.05 mm and with 12-bit resolution.

While the regional registration technique can be used for determining a corresponding structure or region for any structure (both normal tissues and masses) in the breast, in this study we are analyzing its accuracy on biopsy-proven masses alone. The location of the mass on the current mammogram is identified by an Mammography Quality Standards Act (MQSA)-approved radiologist experienced in breast imaging using an interactive image analysis tool on a UNIX workstation. To provide the ground truth for evaluation of the computerized method, the radiologist manually identifies the corresponding region on the previous mammogram. Bounding polygons enclosing the microcalcification cluster on the current mammogram and the corresponding object on the previous mammogram are provided by the radiologist for each case. Each microcalcification cluster as well as the corresponding structure on the previous mammogram are rated for its visibility on a scale of 1 to 10, where the rating of 1 corresponded to the most visible category. The size of the microcalcification cluster on the current mammogram as well as the size of the corresponding structure on the prior mammogram are also measured by the radiologist. The parenchymal density is rated based on the Breast Imaging Reporting and Data System (BI-RADS) lexicon.

(B) Development of a regional registration technique for localization of a search region for the corresponding microcalcification cluster on the prior mammogram of the same view. (Task 2)

We started the development of a multistage regional registration technique for identifying corresponding microcalcification clusters on temporal pairs of mammograms. This detection approach mimics the method used by radiologists for searching corresponding lesions on mammograms, i.e., the lesion is searched at approximately the same radial distance from the nipple on both views, and feature comparison will be used for further identifying the matching lesion. In the first stage, an initial search region was estimated on the prior mammogram based on the lesion location on the current mammogram. In the second stage the search region was refined. In the third stage the lesion was detected within the search region.

Initially, the breast image was segmented from the background in the current and prior mammograms. We used the methods already developed in out lab, which work reliably for segmentation of the breast image from the background for our automated detection algorithms for single images [1], [2].

For the first stage of the multistage regional registration technique we need the nipple location on the current and prior mammograms. We are in the process of developing an automated nipple detection program. Currently its accuracy is about 88% in a data set of about

300 images [3]. However, at this time we used manually marked nipple locations on the mammograms. We are working to further improve the accuracy of the nipple detection algorithm aiming its use into the initial step of our automated interval change analysis scheme.

Initial global alignment of mammograms

In the first stage of registration, an initial fan-shaped search region is automatically defined on the prior mammogram based on the cluster location on the current mammogram. The cluster on the current mammogram can either be detected by an automated program or selected interactively by a radiologist. Currently we used the markings of the cluster locations given by the radiologist .

For the initial estimation of the lesion centroid location on the prior mammogram, we previously developed a regional registration method (RRM) [4][5], based on the radial distance between the nipple and the lesion centroid and the angular distance between the nipple-lesion centroid axis and the breast boundary on the current mammogram.

In this study, the differences in the breast images on the current and prior mammograms were approximately accounted for by warping the current mammogram. We compared both linear and nonlinear global warping of the current mammogram, and studied if these different approaches will improve the localization. In addition we compared the warping techniques to RRM for the initial estimation of the lesion location.

Using the nipple location on both mammograms as an anchor point, the affine (AF) or thin plate splines (TPS) transformation in combination with simplex optimization [6], [7] iteratively warped the current mammogram. The iteration was driven by the maximization of the similarity measure between the breast structures on the current and prior mammograms. Correlation and mutual information (MI) similarity measures were evaluated. An affine transformation [8] is a linear transformation combining scaling, rotation and translation. The transformed object is linearly resized and rotated. In general, the breast images on the two mammograms are distorted with respect to each other nonlinearly. We therefore also investigated if nonlinear warping functions such as thin-plate splines [9], [10] will provide better alignment between the two images. Thin-plate splines warping based on maximization of mutual information have been used successfully for 3D warping of medical images [11]. We studied if this approach will be effective for 2D alignment of breast images.

A set of 390 temporal pairs of mammograms containing biopsy-proven microcalcification clusters or masses was used. 72 temporal pairs containing microcalcification clusters were used for training the parameters of the warping techniques. The remaining 318 pairs were used for testing the performance of the 5 methods. The registration accuracy was analyzed by evaluating the average distance between the centroids of the estimated and the true lesion locations on the prior mammogram.

We found that the average distance between the estimated and the true lesion centroids on the prior mammogram after the initial stage was: RRM = 8.5±6.2mm, correlation-AF = 9.0±6.7 mm, correlation-TPS = 10.3±8.2 mm, MI-AF = 9.2±7.5 mm, MI-TPS = 9.5±8.6 mm.
The RRM method outperformed the warping techniques. It localized the corresponding lesions on temporal pairs of mammograms with the highest accuracy and the lowest standard deviation among the 5 methods.

We will present the preliminary results RSNA , 2003 [12]. We will continue our studies to improve the technique and evaluate its accuracy on a larger data set.

Definition of search region

The location of the cluster on the current mammogram is defined in a polar coordinate system with the nipple as the origin. The position of the microcalcification cluster on the prior mammogram is predicted in a similar manner. An initial fan-shaped search region centered at the predicted location from previous stage of the cluster centroid is then defined on the prior mammogram.

Using a search region with an average area of 1374 mm^2 allowed all clusters for the 72 pairs to be within the fan shape search region. That size of the search region was defined before from the mass local registration [6], which was large enough to include all of the clusters.

We will continue to improve the registration methods in order to refine the localization and reduce the size of the search region.

(C) Adaptation of the automated detection method for identification of candidates of microcalcification clusters within the search region. (Task 3)

The search region (ROI) estimates the area that the cluster is most likely located but it does not provide the exact location. As the next step, automated detection of microcalcification cluster within the search region is performed. Our current automated microcalcification detection algorithm [13] provides a basis for this search. Since the detection is limited to the small search region, the algorithm parameters can be adjusted to improve the detection sensitivity of the very subtle clusters on the prior mammograms without excessive trade-off in increasing false-positives (FPs).

In the second stage of the registration technique, we started investigating the possibility of detection of cluster candidates within the search region with our automated cluster search program with increased sensitivity. At this stage of the project a data set of 72 temporal pairs of mammograms from 31 patients (11 malignant 20 benign) containing biopsy-proven microcalcification clusters was used. The true cluster locations were identified by an MQSA radiologist. Using a search region with an average area of 1374 mm^2 allowed all clusters of interest to be localized in the search region. The average distance between the estimated and the true centroid of the microcalcification clusters on the prior mammogram was 7.9 ± 4.1 mm after the first stage.

Using our current cluster detection program with standard thresholds, 69.4% (50/72) of the clusters (TP) with an average of 0.21 false positives (FP) were detected within the search region on the prior mammogram. Using a high-sensitivity threshold, the sensitivity was increased to 84.7% (61/72) with an average of 0.75 FP within the search region on the prior mammogram.

We will continue to investigate the possibilities to increse the sensitivity without increasing substantially the FP within the search region in order to detect more very subtle clusters.

(D) Feature extraction techniques and initial definition of similarity measure for matching corresponding microcalcification clusters on current and prior mammograms (Task 4)

The cluster (TP) on the current image was paired with every detected cluster (TP or FP) in the search region. Texture and morphological features were extracted from the clusters on the current and the prior mammograms. Difference similarity measure was derived from the extracted features of the TP or FP clusters for each temporal pair. Five difference texture features were obtained.

We will continue the design of new types of features and similarity measures.

(E) Initial design of a correspondence classifier for identification of matched cluster pairs (Task 5)

We started investigation of the ways to design a correspondence classifier. Here we report the initial results.

In the final stage, a correspondence classifier was designed to reduce the false pairs (TP-FP) within the search region. A leave-one-case-out training and testing resampling scheme was used for feature selection and classification.

We used a linear discriminant classifier to merge the selected features for classification of the TP-TP and TP-FP cluster pairs. The correspondence classifier reduced the FP rate to an average of 0.44 FP cluster with sensitivity of 84.7% (61/72).

In the future years we will continue studying and developing different classifiers and ways to represent the correspondence information between prior and current TP and FP clusters.

(F) Initial development of feature measures for characterization of temporal changes in microcalcification clusters. (Task 6)

In this study, a new classification scheme using interval change information was developed to classify mammographic microcalcification clusters as malignant and benign.

From each cluster, 20 run length statistic texture features (RLSF) and 21 morphological features were extracted [14]. Twenty difference RLSF were obtained by subtracting a prior RLSF from the corresponding current RLSF. The feature space consisted of the current and the difference RLSF, as well as the current and the difference morphological features.

In the future years we will develop different classification approaches in order to classify the clusters on malignant and benign.

(6) Key research accomplishments in current year as a result of this grant

- Database collection and extraction of regions of interest (Task 1).
- Further development of methods for establishing corresponding locations in current and previous mammograms (Task 2).
- Starting of the adaptation of the automated detection method for identification of candidates of microcalcification clusters within the search region (Task 3).

- Feature extraction techniques and initial definition of similarity measure for matching corresponding microcalcification clusters on current and prior mammograms (Task 4).
- Initial design of a correspondence classifier for identification of matched cluster pairs (Task 5).
- Initial development of feature measures for characterization of temporal changes in microcalcification clusters (Task 6).

(7) Reportable Outcomes

Publications in current year as a result of this grant

[1] L. Hadjiiski, H.P. Chan, B. Sahiner, C Zhou, M.A. Helvie, M.A. Roubidoux, "Computerized Regional Registration of Corresponding Masses and Microcalcification Clusters on Temporal Pairs of Mammograms for Interval Change Analysis", To be presented at the *89th Scientific Assembly and Annual Meeting of the Radiological Society of North America (RSNA)*, Chicago, Illinois, 2003.

[2] L. Hadjiiski, H.P. Chan, N. Petrick, B. Sahiner, M. Gurcan, M.A. Helvie, at al, "Computerized Regional Registration of Corresponding Microcalcification Clusters on Temporal Pairs of Mammograms for Interval Change Analysis", Presented at the *87th Scientific Assembly and Annual Meeting of the Radiological Society of North America (RSNA)*, Chicago, Illinois, November 25 - 30, 2001. *Radiology* 2001; 221 (P): 425.

[3] L. Hadjiiski, H.P. Chan, M. Gurcan, B. Sahiner, N. Petrick, M.A. Helvie, M. Roubidoux "Computer-Aided Characterization of Malignant and Benign Microcalcification Clusters Based on the Analysis of Temporal Change of Mammographic Features", Presented at the *SPIE International Symposium on Medical Imaging*, San Diego, California, February 23-28, 2002. *Proc. SPIE Medical Imaging, 2002, 4684, pp.749-753..*

Copies of publications are enclosed with this report.

(8) Conclusion

During this year, we have started the development of the regional registration technique. The differences in the breast images on the current and prior mammograms were approximately accounted for by warping the current mammogram. We compared both linear and nonlinear global warping of the current mammogram, and studied if these different approaches will improve the localization. In addition we compared the warping techniques to RRM for the initial estimation of the lesion location. A set of 390 temporal pairs of mammograms containing biopsy-proven microcalcification clusters or masses was used. 72 temporal pairs containing microcalcification clusters were used for training the parameters of the warping techniques. The remaining 318 pairs were used for testing the performance of the 5 methods. The registration accuracy was analyzed by evaluating the average distance between the centroids of the estimated

and the true lesion locations on the prior mammogram. We found that the average distance between the estimated and the true lesion centroids on the previous mammogram after the initial stage was: RRM = 8.5±6.2mm, correlation-AF = 9.0±6.7 mm, correlation-TPS = 10.3±8.2 mm, MI-AF = 9.2±7.5 mm, MI-TPS = 9.5±8.6 mm. The RRM method outperformed the warping techniques. It localized the corresponding lesions on temporal pairs of mammograms with the highest accuracy and the lowest standard deviation among the 5 methods.

In the second stage of the registration technique, we started investigating the possibility of detection of cluster candidates within the search region with an automated cluster search program with increased sensitivity. At this stage of the project a data set of 72 temporal pairs of mammograms from 31 patients (11 malignant 20 benign) containing biopsy-proven microcalcification clusters was used. The true cluster locations were identified by an MQSA radiologist. A leave-one-case-out training and testing resampling scheme was used for feature selection and classification. Using a search region with an average area of 1374 mm^2 allowed all clusters of interest to be localized in the search region. The average distance between the estimated and the true centroid of the microcalcification clusters on the prior mammogram was 7.9 ± 4.1 mm after the first stage. Using our current cluster detection program with standard thresholds, 69.4% (50/72) of the clusters (TP) with an average of 0.21 false positives (FP) were detected within the search region on the prior mammogram. Using a high-sensitivity threshold, the sensitivity was increased to 84.7% (61/72) with an average of 0.75 FP within the search region on the prior mammogram.

The cluster (TP) on the current image was paired with every detected cluster (TP or FP) in the search region. Texture and morphological features were extracted from the clusters on the current and the prior mammograms. Difference similarity measure was derived from the extracted features of the TP or FP clusters for each temporal pair.

We used a linear discriminant classifier to merge the selected features for classification of the TP-TP and TP-FP cluster pairs. The correspondence classifier reduced the FP rate to an average of 0.44 FP cluster with sensitivity of 84.7% (61/72).

In this study, a new classification scheme using interval change information was developed to classify mammographic microcalcification clusters as malignant and benign. From each cluster, 20 run length statistic texture features (RLSF) and 21 morphological features were extracted. Twenty difference RLSF were obtained by subtracting a prior RLSF from the corresponding current RLSF. The feature space consisted of the current and the difference RLSF, as well as the current and the difference morphological features.

(9) References

1. Petrick N, Chan HP, Wei D, Sahiner B, Helvie MA and Adler DD, "Automated detection of breast masses on mammograms using adaptive contrast enhancement and texture classification," Med Phys 23, 1685-1696 (1996).
2. Zhou C, Chan H, Petrick N, Helvie M, Goodsitt M, Sahiner B and Hadjiiski L, "Computerized image analysis: Estimation of breast density on mammograms.," Med Phys.,28 (6), June (2001) pp. 1056-1069.
3. Zhou C, Chan HP, Petrick N, Goodsitt MM, Paramagul C and Hadjiiski LM, "Computerized image analysis: breast segmentation and nipple identification on mammograms," Chicago

2000-World Congress on Medical Physics and Biomedical Engineering., Chicago, Illinois, July 23-28, 2000,

4. S.S. Gopal, H.P. Chan, T.E. Wilson, M.A. Helvie, N. Petrick, B. Sahiner, "A regional registration technique for automated interval change analysis of breast lesions on mammograms", *Medical Physics,* 1999, 26:2669-2679.
5. L. Hadjiiski, H.P. Chan, B. Sahiner, N. Petrick, M. Helvie, "Automated Registration of Breast Lesions in Temporal Pairs of Mammograms for Interval Change Analysis – Local Affine Transformation for Improved Localization", *Medical Physics*, 28 (6), June 2001, pp. 1070-1079.
6. Rao SS, *Optimization: Theory and Applications*, (Wiley Eastern Limited, 1979).
7. Lootsma FA, *Numerical methods for non-linear optimization*, (Academic Press, New York, 1972).
8. Quan L and Kanade T, "Affine structure from line correspondance with uncalibrated affine cameras," IEEE Trans Pat Anal Machine Intel 19, 834-845 (1997).
9. Bookstein LF, *Morphometric Tools for Landmark Data*, (Cambridge University Press, Cambridge, 1991).
10. Bookstein FL and Green WDKA, "A feature spase for edgels in images with landmarks," J Math Imag Vision 3, 231-261 (1993).
11. Meyer CR, Boes JL, Kim B, Bland P, Zasadny KR, Kison PC, Koral K, Frey KA and Wahl RL, "Demonstration of accuracy and clinical versatility of mutual information for automatic multimodality image fusion using affine and thin plate spline warped geometric deformations," Medical Image Analysis 3, 195-206 (1997).
12. L. Hadjiiski, H.P. Chan, B. Sahiner, C Zhou, M.A. Helvie, M.A. Roubidoux, "Computerized Regional Registration of Corresponding Masses and Microcalcification Clusters on Temporal Pairs of Mammograms for Interval Change Analysis", To be presented at the *89th Scientific Assembly and Annual Meeting of the Radiological Society of North America (RSNA)*, Chicago, Illinois, 2003.
13. Chan HP, Lo SCB, Sahiner B, Lam KL and Helvie MA, "Computer-aided detection of mammographic microcalcifications: Pattern recognition with an artificial neural network," Med Phys 22, 1555-1567 (1995).
14. L. Hadjiiski, H.P. Chan, M. Gurcan, B. Sahiner, N. Petrick, M.A. Helvie, M. Roubidoux "Computer-Aided Characterization of Malignant and Benign Microcalcification Clusters Based on the Analysis of Temporal Change of Mammographic Features", Presented at the *SPIE International Symposium on Medical Imaging*, San Diego, California, February 23-28, 2002. *Proc. SPIE Medical Imaging, 2002, 4684, pp.749-753.*

(10) Appendix

Copies of publications are enclosed with this report.

Accepted For Presentation at RSNA 2003

Computerized Regional Registration of Corresponding Masses and Microcalcification Clusters on Temporal Pairs of Mammograms for Interval Change Analysis

Lubomir Hadjiiski, Heang-Ping Chan, Berkman Sahiner, Chuan Zhou, Mark A. Helvie, Marilyn Roubidoux

PURPOSE: To develop a registration technique for automated identification of corresponding lesions on a temporal pair of mammograms of the same view. This technique is the basis for interval change analysis of breast lesions in CAD applications.

MATERIALS AND METHODS: A multi-stage registration technique is being developed. In the first stage, an initial search region was estimated on the prior mammogram based on the lesion location on the current mammogram. In the second stage the search region was refined. In the third stage the lesion was detected within the search region.

For the initial estimation of the lesion centroid location on the prior mammogram, we previously developed a regional registration method (RRM), based on the radial distance between the nipple and the lesion centroid and the angular distance between the nipple-lesion centroid axis and the breast boundary on the current mammogram. In the present study, we compared the RRM to the use of warping techniques for the initial estimation of the lesion location. The current mammogram was warped by affine (AF) or thin plate splines (TPS) transformation in combination with simplex optimization in order to maximize a similarity measure between the breast areas on the current and prior mammograms. Correlation and mutual information (MI) similarity measures were evaluated.

A set of 390 temporal pairs of mammograms containing biopsy-proven masses or microcalcification clusters was used. The true lesion locations were identified by an MQSA radiologist on all mammograms. 72 temporal pairs were used for training the parameters of the warping techniques. The remaining 318 pairs were used for testing the performance of the 5 methods. The registration accuracy was analyzed by evaluating the average distance between the centroids of the estimated and the true lesion locations on the prior mammogram.

RESULTS: The average distance between the estimated and the true lesion centroids on the previous mammogram after the initial stage was: RRM = 8.5±6.2mm, correlation-AF = 9.0±6.7 mm, correlation-TPS = 10.3±8.2 mm, MI-AF = 9.2±7.5 mm, MI-TPS = 9.5±8.6 mm. After the final registration stage, the average distance between the estimated and the true centroids was: RRM=6.4±8.9mm, correlation-AF = 7.0±9.5 mm, correlation-TPS = 7.4±10.2 mm, MI-AF = 6.9±9.5 mm, MI-TPS = 7.2±11.1 mm.

CONCLUSIONS: The RRM method outperformed the warping techniques. It localized the corresponding lesions on temporal pairs of mammograms with the highest accuracy and the lowest standard deviation among the 5 methods.

RSNA 2001

SUPPLEMENT TO RADIOLOGY • NOVEMBER 2001 • VOLUME 221 (P)

PERSONAL PLANNER, PAGE 981

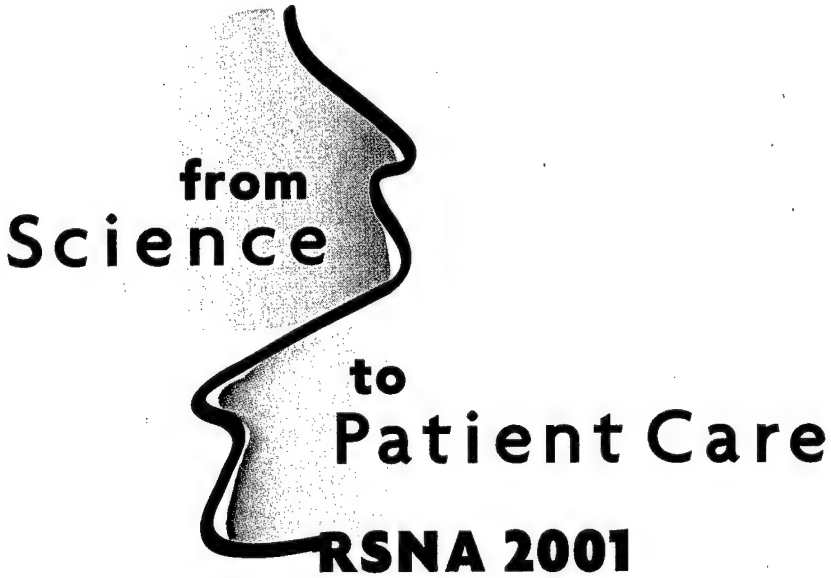
SCIENTIFIC PROGRAM

THE RADIOLOGICAL SOCIETY OF NORTH AMERICA

87th SCIENTIFIC ASSEMBLY AND ANNUAL MEETING

NOVEMBER 25 — 30, 2001

McCORMICK PLACE, CHICAGO, ILLINOIS



815 • 2:48PM

Computerized Regional Registration of Corresponding Microcalcifica Clusters on Temporal Pairs of Mammograms for Interval Change Analy

L.M. Hadjiiski, PhD, Ann Arbor, MI (lhadjisk@umich.edu) • H. Chan, PhD • Petrick, PhD • B. Sahiner, PhD • M.N. Gurcan, PhD • M.A. Helvie, MD • et al

PURPOSE: To develop a regional registration technique for identify corresponding microcalcification clusters on current and prior mam grams of the same view. The technique will be useful for computeri analysis of interval changes of microcalcification clusters in comp aided diagnosis (CAD).

METHOD AND MATERIALS: A multi-stage regional registration t nique is being developed. In the first stage, an initial fan-shape se region was estimated on the prior mammogram based on the clu location on the current mammogram. In the second stage, detectio cluster candidates within the search region was performed with automated cluster search program. The cluster (TP) on the current im was paired with every detected cluster (TP or FP) in the search regio the final stage, a correspondence classifier was designed to reduce the f pairs (TP-FP) within the search region. Texture and morphological feat were extracted from the clusters on the current and the prior mam grams. Similarity measures were derived from the extracted features o TP or FP clusters for each temporal pair. Stepwise feature selection simplex optimization was used to select the optimal feature subse linear discriminant classifier was used to merge the selected feature classification of the TP-TP and TP-FP cluster pairs. In this prelimi study, a data set of 51 temporal pairs of mammograms from 19 pati containing biopsy-proven microcalcification clusters was used. The cluster locations were identified by an MQSA radiologist. A leave-one- out training and testing resampling scheme was used for feature selec and classification.

RESULTS: Using a search region with an average area of 1350 allowed all clusters of interest to be localized in the search region. average distance between the estimated and the true centroid of microcalcification clusters on the prior mammogram was 7.9±4.1 mm the first stage. The cluster search program detected 90% (46/51) of the clusters with an average of 0.69 FP cluster within the search region o prior mammograms. The correspondence classifier reduced the FP ra an average of 0.41 FP cluster at the cost of misclassifying 1 true pair.

CONCLUSIONS: Our preliminary study demonstrated that the regi registration technique is a promising approach for identifying corresp ing microcalcification clusters on temporal pairs of mammograms. Fu studies are underway to improve the technique and to evaluat accuracy on a larger data set.

Computer-aided characterization of malignant and benign microcalcification clusters based on the analysis of temporal change of mammographic features

Lubomir Hadjiiski, Heang-Ping Chan, Berkman Sahiner, Nicholas Petrick,
Mark A. Helvie, Marilyn Roubidoux, Metin Gurcan

Department of Radiology, The University of Michigan, Ann Arbor, MI 48109-0904

ABSTRACT

We have previously demonstrated that interval change analysis can improve differentiation of malignant and benign masses. In this study, a new classification scheme using interval change information was developed to classify mammographic microcalcification clusters as malignant and benign. From each cluster, 20 run length statistic texture features (RLSF) and 21 morphological features were extracted. Twenty difference RLSF were obtained by subtracting a prior RLSF from the corresponding current RLSF. The feature space consisted of the current and the difference RLSF, as well as the current and the difference morphological features. A leave-one-case-out resampling was used to train and test the classifier using 65 temporal image pairs (19 malignant, 46 benign) containing biopsy-proven microcalcification clusters. Stepwise feature selection and a linear discriminant classifier, designed with the training subsets alone, were used to select and merge the most useful features. An average of 12 features were selected from the training subsets, of which 3 difference RLSF and 7 morphological features were consistently selected from most of the training subsets. The classifier achieved an average training A_z of 0.98 and a test A_z of 0.87. For comparison, a classifier based on the current single image features achieved an average training A_z of 0.88 and test A_z of 0.81. These results indicate that the use of temporal information improved the accuracy of microcalcification characterization.

Keywords: Computer-Aided Diagnosis, Interval Changes, Classification, Feature analysis, Mammography, Malignancy.

1. INTRODUCTION

Mammography is currently the most effective method for early breast cancer detection[1,2]. Radiologists routinely compare mammograms from a current examination with those obtained in previous years, if available, for identifying interval changes, detecting potential abnormalities, and evaluating breast lesions. It is widely accepted that analysis of interval changes in mammographic features is very useful for both detection and classification of abnormalities[3,4]. A variety of computer-aided diagnosis (CAD) techniques have been developed to detect mammographic abnormalities and to distinguish between malignant and benign lesions. We are studying the use of CAD techniques to assist radiologists in interval change analysis.

Commonly used classification methods for CAD use information from a single image. These methods have been shown to perform well in lesion classification problems[5-11]. However, when multiple-year mammograms of a lesion are available, it is not trivial to design computer vision methods to use the temporal information for computer-aided classification and to improve the differentiation between benign and malignant masses.

The goal of our research is to develop a technique for computerized analysis of temporal differences between a microcalcification cluster on the most recent mammogram and a prior mammogram of the same view. The computer algorithm can be used to assist radiologists in evaluating interval changes and thus distinguishing between malignant and benign microcalcification clusters for CAD. In our previous studies we have demonstrated that interval change analysis can improve differentiation of malignant and benign masses[12,13]. In this study we will introduce a new classification scheme using interval change information to classify mammographic microcalcification clusters as malignant and benign. Additionally, we will compare this method with a classification method based on information extracted from the current mammogram alone.

Medical Imaging 2002: Image Processing, Milan Sonka, J. Michael Fitzpatrick,
Editors, Proceedings of SPIE Vol. 4684 (2002) © 2002 SPIE · 1605-7422/02/$15.00

2. CLASSIFICATION TECHNIQUE

A new classification scheme was developed to classify mammographic microcalcification clusters as malignant and benign by using interval change information. The technique is based on the design of features that will represent the temporal information and will discriminate between malignant and benign microcalcification clusters.

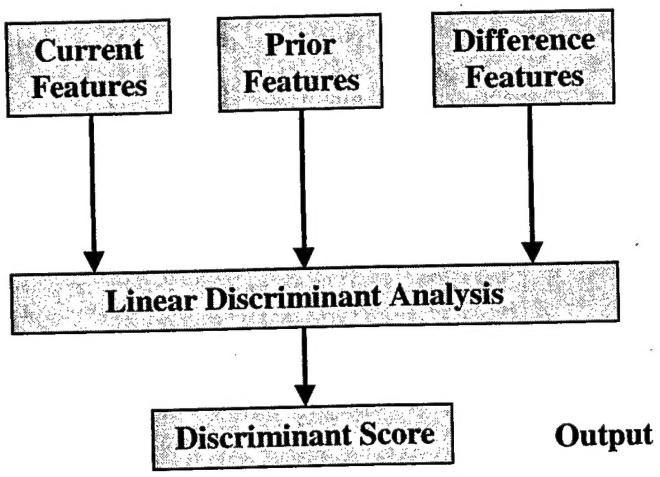


Figure 1. Block-diagram of the classification method.

The clusters to be analyzed can either be identified manually by a radiologist or automatically by a computerized detection program. In this study, the microcalcification clusters were identified by an MQSA radiologist on each mammogram. The locations of the individual microcalcifications from the clusters on both the current and the prior mammograms have been determined manually. Features such as texture features, morphological features and the number of microclacifications in a cluster were extracted from each microcalcification cluster. Additionally, the interval change of a given feature of the cluster is determined as the difference between its current feature value and the corresponding prior feature value. The feature space consisted of current, prior, and difference features. Stepwise feature selection applied to linear discriminant analysis (LDA) was used to select the most useful features. The selected features were then used as the input predictor variables of the LDA classifier (Figure 1). A leave-one-case-out resampling scheme was employed to train and test the classifier.

To evaluate the improvement in the classifier performance designed by using the temporal change information, an additional classifier was trained using the information extracted from the current images of the temporal pairs. We will refer to these images as current images. Comparison of the two classifiers will reveal the effectiveness of interval change analysis on classification of malignant and benign microcalcification clusters.

3. DATA SET

In this preliminary study, 65 temporal image pairs from 29 patients containing biopsy-proven microcalcification clusters on the current mammograms were chosen from patient files. Eleven of the cases were malignant and 18 were benign. For the 29 patients 102 mammograms were chosen. The mammograms were digitized with a LUMISCAN 85 laser scanner at a pixel resolution of 50 μm × 50 μm and 4096 gray levels. The digitizer was calibrated so that gray level values were linearly proportional to the optical density (OD) within the range of 0 to 4 OD units, with a slope of 0.001 OD/pixel value. Outside this range, the slope of the calibration curve decreased gradually. The digitizer output was linearly converted so that a large pixel value corresponded to a low optical density. The images were averaged and down-sampled by a factor of 2 resulting in images with a pixel size of 100 μm × 100 μm for further analysis.

The 102 mammograms contained different mammographic views and multiple years of the microcalcification clusters including the year when the biopsy was performed. By matching microcalcification clusters of the same view from two different exams, a total of 65 temporal pairs were formed, of which 19 were malignant and 46 benign. A malignant temporal pair consisted of a biopsy proven malignant microcalcification cluster or a cluster that was followed up and found to be

malignant by biopsy in a future year. Similar definitions were used for the benign temporal pairs. Within the 65 temporal pairs, a total of 56 mammograms were single current mammograms. Of the 56 current mammograms, 16 were malignant and 40 benign. Since all cases in this data set had undergone biopsy, the benign clusters in this set could not be distinguished easily from malignant ones based on current image criteria.

For the malignant microcalcification clusters in this data set, the average cluster size was 8.8 mm on the prior mammograms and 15.1 mm on the current mammograms. The corresponding sizes were 11.4 mm and 11.6 mm, respectively, for the benign microcalcification clusters. The temporal pairs had a time interval of 3 to 32 months. Approximately 50% of the pairs had a time interval of 6 months and more than 30% had a time interval of 12 months.

4. FEATURE EXTRACTION

A rectangular region of interest (ROI) was defined to include the radiologist-identified microcalcification cluster with an additional surrounding breast tissue region of at least 40 pixels wide from any point of the cluster boundary.

The texture features used in this study were calculated from run-length statistics (RLS) matrices[14]. The RLS matrices were computed from the defined ROIs. RLS texture features were extracted from the vertical and horizontal gradient magnitude images, which were obtained by filtering the ROI image with horizontally or vertically oriented Sobel filters and computing the absolute gradient value of the filtered image. Five texture measures, namely, short run emphasis, long run emphasis, gray level nonuniformity, run length nonuniformity, and run percentage were extracted from the vertical and horizontal gradient images in two directions, $\theta = 0^\circ$, and $\theta = 90^\circ$. Therefore, a total of 20 RLS features were calculated for each ROI. The definition of the RLS feature measures can be found in the literature[14].

For the extraction of the morphological features, the locations of the individual microcalcifications in a cluster were identified manually. The true microcalcification were defined as those visible on the film mammograms with a magnifier. The morphological features included features describing the variations of the shape and size of the individual microcalcifications in a cluster such as the area, mean density, eccentricity, moment ratio and axis ratio[11]. To quantify the variation of the visibility and shape descriptors in a cluster, the maximum, the average, the standard deviation and the coefficient of variation were calculated for each feature. The number of microcalcifications in a cluster[11] was also included as a feature.

A total of 41 features (20 RLS and 21 morphological) were extracted from each microcalcification cluster. Additionally, difference features were obtained by subtracting a prior feature from the corresponding current feature. Therefore 20 RLS and 21 morphological difference features were obtained.

5. FEATURE SELECTION

In order to reduce the number of the features and to obtain the best feature subset to design an effective classifier, feature selection with stepwise linear discriminant analysis[15,16] was applied. At each step of the stepwise selection procedure, one feature is entered or removed from the feature pool based on analysis of its effect on the selection criterion. The stepwise selection procedure is controlled by a simplex optimization method[17,18] in such a way that a minimum number of features were selected to achieve a high accuracy of classification by LDA. More details about the stepwise linear discriminant analysis and its application to CAD can be found elsewhere[5].

6. EVALUATION METHODS

To evaluate the classifier performance, the training and test discriminant scores were analyzed using receiver operating characteristic (ROC) methodology[19]. The discriminant scores of the malignant and benign masses were used as decision variables in the LABROC1 program[20], which fits a binormal ROC curve based on maximum likelihood estimation. The classification accuracy was evaluated as the area under the ROC curve, A_z. The performances of the classifiers were also assessed by estimation of the partial area index ($A_z^{(0.9)}$). $A_z^{(0.9)}$ is defined as the area that lies under the ROC curve but above a sensitivity threshold of 0.9 ($TPF_0 = 0.9$) normalized to the total area above TPF_0, (1-TPF_0). It indicates the performance of the classifier in the high sensitivity (low false negative) region that is most important for a cancer detection task.

7. CLASSIFICATION RESULTS

For the data set used in this study, an average of 12 features were selected from the 29 training subsets. The most frequently selected features included 3 difference RLS features and 9 morphological features from the current image. Three difference RLSF and 7 morphological features were consistently selected from most of the training subsets. The LDA classifier achieved an average training A_z of 0.98 and a test A_z of 0.87. The LDA classifier using features extracted from the current images of the temporal pairs achieved an average training A_z of 0.88 and a test A_z of 0.81. An average of 4 features were selected from the 29 training subsets. The most frequently selected features were 1 RLS feature and 3 morphological features.

The difference in the test A_z between the two classifiers did not achieve statistical significance. The classifier based on temporal pairs achieved a test partial $A_z^{(0.9)}$ of 0.63 and the classifier based on current images achieved a test $A_z^{(0.9)}$ of 0.43. These results are summarized in Table 1.

Table 1. Classification accuracy for the classifier based on the temporal change information and the classifier based on current single image information.

Classification	Avg. no. Of selected features	Training A_z	Test A_z	Test partial $A_z^{(0.9)}$
Temporal pairs	12	0.98	0.87 ± 0.044	0.63
Current images	4	0.88	0.81 ± 0.059	0.43

8. CONCLUSION

The difference RLS texture features and the current morphological features were useful for identification of malignant microcalcifications in temporal pairs of mammograms. The information on the prior image was important for characterization of the microcalcifications; 3 out of the 12 selected features contained prior information. Morphological features describing the variations of the shape and size of the individual microcalcifications in a cluster were more effective than the features related to the cluster size. This is probably due to the fact that we used biopsy-proven cases in this study and many of the biopsied benign microcalcification clusters also grew over time. Combination of current and temporal change information improved classification accuracy compared to current information alone in terms of the A_z. The increase in A_z (0.06), although large, did not achieve statistical significance due to the small sample size. The partial area under the ROC curve is also improved for the classifier based on current and prior images ($A_z^{(0.9)} = 0.63$) compared to the classifier based only on the current images ($A_z^{(0.9)} = 0.43$). However, the difference did not achieve statistical significance either. Further studies are underway to improve this temporal change classification technique and to evaluate its performance on a larger data set.

ACKNOWLEDGMENTS

This work is supported by a Career Development Award from the USAMRMC (DAMD 17-98-1-8211) (L.H.), a Basic Radiological Science Innovative Research Award, 2001 (L.H.), from the University of Michigan, and a USPHS Grant CA 48129. The content of this publication does not necessarily reflect the position of the funding agency, and no official endorsement of any equipment and product of any companies mentioned in this publication should be inferred.

REFERENCES

1. H. C. Zuckerman, "The role of mammography in the diagnosis of breast cancer," *In: Breast Cancer, Diagnosis and Treatment*, 152-172, Eds. I. M. Ariel and J. B. Cleary, McGraw-Hill, New York, 1987.

2. L. Tabar and P. B. Dean, "The Control of Breast Cancer through Mammography Screening," *Radiologic Clincs of North America* **25**, 961, 1987.

3. L. W. Bassett, B. Shayestehfar, and I. Hirbawi, "Obtaining previous mammograms for comparison: usefullness and costs," *Amer. J. Roentgenology* **163**, 1083-1086, 1994.

4. E. A. Sickles, "Periodic mammographic follow-up of probably benign lesions: results in 3183 consecutive cases," *Radiology* **179**, 463-468, 1991.

5. B. Sahiner, H. P. Chan, N. Petrick, M. A. Helvie, and M. M. Goodsitt, "Computerized characterization of masses on mammograms: The rubber band straightening transform and texture analysis," *Medical Physics* **25**, 516-526, 1998.

6. Z. M. Huo, M. L. Giger, C. J. Vyborny, D. E. Wolverton, R. A. Schmidt, and K. Doi, "Automated computerized classification of malignant and benign masses on digitized mammograms," *Academic Radiology* **5**, 155-168, 1998.

7. L. M. Hadjiiski, B. Sahiner, H. P. Chan, N. Petrick, and M. A. Helvie, "Classification of malignant and benign masses based on hybrid ART2LDA approach," *IEEE Transactions on Medical Imaging* **18**, 1178-1187, 1999.

8. H. P. Chan, D. Wei, K. L. Lam, B. Sahiner, M. A. Helvie, D. D. Adler, and M. M. Goodsitt, "Classification of malignant and benign microcalcifications by texture analysis," *Medical Physics* **22**, 938, 1995.

9. Y. Jiang, R. M. Nishikawa, D. E. Wolverton, C. E. Metz, M. L. Giger, R. A. Schmidt, C. J. Vyborny, and K. Doi, "Malignant and benign clustered microcalcifications: automated feature analysis and classification," *Radiology* **198**, 671-678, 1996.

10. H. P. Chan, B. Sahiner, N. Petrick, M. A. Helvie, K. L. Leung, D. D. Adler, and M. M. Goodsitt, "Computerized classification of malignant and benign microcalcifications on mammograms: texture analysis using an artificial neural network," *Physics in Medicine and Biology* **42**, 549-567, 1997.

11. H. P. Chan, B. Sahiner, K. L. Lam, N. Petrick, M. A. Helvie, M. M. Goodsitt, and D. D. Adler, "Computerized analysis of mammographic microcalcifications in morphological and texture feature space," *Medical Physics* **25**, 2007-2019, 1998.

12. L. M. Hadjiiski, B. Sahiner, H. P. Chan, N. Petrick, M. A. Helvie, and M. Gurcan, "Analysis of temporal change of mammographic features for computer-aided characterization of malignant and benign masses," *Proceedings of the SPIE* **4322**, 661-666, 2001.

13. L. M. Hadjiiski, B. Sahiner, H. P. Chan, N. Petrick, M. A. Helvie, and M. N. Gurcan, "Analysis of Temporal Change of Mammographic Features: Computer-Aided Classification of Malignant and Benign Breast Masses," *Medical Physics* **28**, 2309-2317, 2001.

14. M. M. Galloway, "Texture classification using gray level run lengths," *Computer Graphics and Image Processing* **4**, 172-179, 1975.

15. M. J. Norusis, *SPSS for Windows Release 6 Professional Statistics*, SPSS Inc., Chicago, IL, 1993.

16. M. M. Tatsuoka, *Multivariate Analysis, Techniques for Educational and Psychological Research*, 2nd ed. Macmillan, New York, 1988.

17. S. S. Rao, *Optimization: Theory and Applications*, Wiley Eastern Limited, 1979.

18. F. A. Lootsma, *Numerical methods for non-linear optimization*, Academic Press, New York, 1972.

19. C. E. Metz, "ROC methodology in radiologic imaging," *Investigative Radiology* **21**, 720-733, 1986.

20. C. E. Metz, J. H. Shen, and B. A. Herman, "New methods for estimating a binormal ROC curve from continuously-distributed test results," *Annual Meeting of the American Statistical Association*, Anaheim, CA, 1990.